

Value matters: Predictability of Stock Index Returns

Natascia Angelini,^a Giacomo Bormetti,^{b,c} Stefano Marmi,^b and Franco Nardini^a

March 3, 2013

^a *Department of Mathematics for the Social Sciences, University of Bologna, Viale Filopanti 5,
40126 Bologna, Italy*

^b *Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy*

^c *National Institute of Nuclear Physics, Pavia Division, Via A. Bassi 6, 27100 Pavia, Italy*

Abstract

The aim of this paper is twofold: to provide a theoretical framework and to give further empirical support to Shiller's test of the appropriateness of prices in the stock market based on the Cyclically Adjusted Price Earnings (CAPE) ratio. We devote the first part of the paper to the empirical analysis and we show that the CAPE is a powerful predictor of future long run performances of the market not only for the U.S. but also for countries such as Belgium, France, Germany, Japan, the Netherlands, Norway, Sweden and Switzerland. We show four relevant empirical facts: i) the striking ability of the logarithmic averaged earning over price ratio to predict returns of the index, with an R squared which increases with the time horizon, ii) how this evidence increases switching from returns to gross returns, iii) moving over different time horizons, the regression coefficients are constant in a statistically robust way, and iv) the poorness of the prediction when the precursor is adjusted with long term interest rate. In the second part we provide a theoretical justification of the empirical observations. Indeed we propose a simple model of the price dynamics in which the return growth depends on three components: a) a momentum component, naturally justified in terms of agents' belief that expected returns are higher in

bullish markets than in bearish ones; b) a fundamental component proportional to the log earnings over price ratio at time zero. The initial value of the ratio determines the reference growth level, from which the actual stock price may deviate as an effect of random external disturbances, and c) a driving component ensuring the diffusive behaviour of stock prices. Under these assumptions, we are able to prove that, if we consider a sufficiently large number of periods, the expected rate of return and the expected gross return are linear in the initial time value of the log earnings over price ratio, and their variance goes to zero with rate of convergence equal to minus one. Ultimately this means that, in our model, the stock prices dynamics may generate bubbles and crashes in the short and medium run, whereas for future long-term returns the valuation ratio remains a good predictor.

JEL codes: G12, G17

Keywords: Valuation Ratios, Long Run Stock Market Returns

1 Introduction

A question that has been asked with a certain regularity since the establishment of the first stock index is the following, see [Myers and Swaminathan \(1999\)](#): is it possible to calculate “the Intrinsic Value of the Dow”? Or, more generally: what is the intrinsic value of Wall Street? A first commonsense answer is that this value may be calculated by summing “the intrinsic value” of each quoted company. At the beginning of the past century, see [Williams \(1938\)](#), [Graham and Meredith \(1998\)](#), it was well known that the latter depends on the expected future earnings and on the balance sheet, but unfortunately there was no easy formula, which yielded even a rough approximation. Therefore aggregating all expectations for all companies quoted in the market soon appeared to be a challenging task.

Between the Fifties and the Sixties the researchers’ attitude towards the problem changed radically. Modern portfolio theory and the statistical evidence that stock prices follow random walks, see [Mandelbrot \(1965\)](#), [Fama \(1965\)](#), has led to the Efficient Market Hypothesis, see [Samuelson \(1965\)](#), [Fama \(1970\)](#). According to this theory, financial markets are “informationally efficient”, that is, one can not achieve returns in excess of average market returns on a risk-adjusted basis, given the information available at the time the investment is made. If this were the case “the intrinsic value” of a stock would be simply its market price and the

problem would ultimately be non-existent. The bubble burst in 1987 and the exceptional price boom in the late '90 lead more and more researcher to cast doubts on the hypothesis' truth. At the very beginning of 2000 Robert Shiller wrote: "we do not know whether the market level makes any sense, or whether they are indeed the result of some human tendency that might be called irrational exuberance", [Shiller \(2000\)](#), which is apparently at odds with the assumption that market prices are always fair prices. Shiller's skeptical attitude was anticipated in one of the most famous books ever written on the stock market: in 1934, Graham and Dodd strongly advocated the fundamental approach to investment valuation and recommended to "shift(s) the original point of departure, or basis of computation, from the current earnings to the average earnings, which should cover a period of not less than five years, and preferably seven to ten years" [Graham et al. \(1962\)](#). Since 1988 Campbell and Shiller have been focusing their attention on average earnings as predictor of future dividends¹ and future stock prices²; [Shiller \(2000\)](#) proposed an innovative test of the appropriateness of prices in the stock market: the Cyclically Adjusted Price Earning ratio, and showed that it is a powerful predictor of future long run performances of the market; the performance of the test is quite satisfactory in the case of the US market from the end of 19th century up to today. In the same years, again elaborating on [Graham et al. \(1962\)](#), [Lander et al. \(1997\)](#) suggested an alternative but similar criterion based on the assumption "that many investors are constantly making a choice between stock and bond purchases; as the yield on bonds advances, they would be expected to demand a correspondingly higher return on stocks, and conversely as bond yields decline [Graham et al. \(1962, pag. 510\)](#)"; their "approach considers whether stocks are appropriately valued relative to analysts' perceptions of future earnings and yields on alternative investments" such as high grade corporate bonds (for a comparison between the two approaches see [Weigand and Irons \(2006\)](#)).

A first goal of this paper is to provide further empirical support to the Shiller test by considering the dependence of future returns on the valuation of the stock index at the time of investing in several countries different from the United States (more precisely we will investigate

¹"Our results indicate that a long moving average of real earnings helps to forecast future real dividends. The ratio of this earnings variable to the current stock price is a powerful predictor of the return on stock, particularly when the return is measured over several years", see [Campbell and Shiller \(1988a\)](#).

²The price-smoothed earnings ratio has little ability to predict future growth in smoothed earnings; the R squared statistics are less than 4% over one year and over ten years. The ratio is a good forecaster of ten-year growth in stock prices, with an R squared statistic of 37%, see [Campbell and Shiller \(1998\)](#).

the relationship for Australia, Belgium, Canada, France, Germany, Japan, the Netherlands, Norway, Sweden, Switzerland and the United Kingdom). For all these countries, even though the historical series are much shorter than those available for the US, we are able to perform an exhaustive analysis of the relationship between CAPE and long-term index returns. The empirical analysis confirms the ability of the CAPE to predict future long run performances of the market also in the case of these countries. The evidence is robust: a regression analysis shows that the logarithmic averaged earning over price ratio predicts the future returns of the index, with an R squared which is steadily increasing the longer the future time horizon is. Indeed this evidence increases when switching from returns to gross returns, and moving over different time horizons, the regression coefficients are constant in a statistically robust way. We also remark how attempting to adjust the precursor with long term interest rate, as suggested by the so-called Fed model, negatively affects the accuracy of the prediction.

A second goal of this paper is to provide a theoretical framework which explains how future stock index returns depend on the Cyclically Adjusted Price Earnings (CAPE) ratio. This is achieved by constructing a model.

What we require of our model is similar to what Chiarella required in [Chiarella \(1992, see pag. 102\)](#)

“Firstly we require that the model generate a significant transitory component around the equilibrium which reflects the rationally expected value of the asset. Secondly the model must allow for the incorporation of chartists, a group which bases its market actions on an analysis of past trends. Since this group seems to be an important part of real markets it is important to determine what effect its activity has on the behavior of asset prices and whether the behavior of a model incorporating chartists comes closer to explaining some of the empirical results cited earlier”.

We choose, however, a different approach and we work at an aggregate level instead of explicitly modelling two different types of agents. We introduce a momentum component in a “simple error-correction model that predicts the return of the S&P based on the deviations from a presumed equilibrium between . . . earnings yields” [Lander et al. \(1997, see pag. 3\)](#) and a long run target yield rate. The return growth depends on three components

- a) a momentum component, naturally justified in terms of agents’ expectation that expected

returns are higher in bullish markets than in bearish ones;

- b) a fundamental component, proportional to a function of the level of the logarithmic averaged earnings over price ratio (for brevity log EP ratio) at time zero; the initial time value of log EP ratio determines the reference growth level, from which the actual stock price may deviate as an effect of random external disturbances;
- c) a driving component ensuring the diffusive behavior of stock prices.

Under these assumptions, we are able to prove that, if we consider a sufficiently large number of periods, the expected rate of return and the expected gross return are linear in the initial time value of log EP, and their variance converges to zero with rate of convergence equal to minus one. This means that, in our model, the stock prices dynamics may generate bubbles and crashes in the short and medium run, nevertheless the averaged earnings price ratio is a good predictor of future long-run returns, as claimed by [Campbell and Shiller \(1988a\)](#), [Shiller \(2000\)](#), [Lander et al. \(1997\)](#).

2 Data set and empirical facts

The data set analyzed in this paper consists of records on a monthly basis of prices, earnings, and dividends for stock price indexes from twelve developed countries. In more detail, for the US market we consider the Standard and Poor Composite Stock Price Index data discussed in [Campbell and Shiller \(1987, 1988a,b\)](#), and freely available from Robert J. Shiller's webpage <http://www.econ.yale.edu/~shiller/>. These time series cover the entire period from January 1871 until March 2011. For Australia, Belgium, Canada, France, Germany, Japan, the Netherlands, Norway, Sweden, Switzerland, and United Kingdom we resort to MSCI Indexes, MSCI Dividend Yields, and MSCI Price over Earning Ratios, kindly provided by Factset³. At variance with US data, these series span a narrower time window running from December 1969 to December 2010. All nominal series are in local currencies and have been deflated using Consumer Price Indexes available at the Federal Reserve Economic Data repository <http://research.stlouisfed.org/fred2/>. In the very early part of our work we also wish to study excess index returns over long term

³<http://www.factset.com/>.

debt for the US market. The interest rate we use is the 10-Year Treasury Constant Maturity Rate (GS10).

The real price of the stock index, measured at the beginning of time period t , will be written P_t , while for the real dividend paid on the portfolio during the period between t and $t + 1$ we will use D_t . Accordingly, the realized log gross return on the index, held from the beginning of time t and the beginning of time $t + 1$ is written

$$H_t = \log(P_{t+1} + D_t) - \log P_t.$$

Since data are recorded every month, we naturally refer the notation $t + 1$ to the time instant t increased by one month. Accordingly, the realized gross yield on a monthly⁴ basis over h periods is

$$y_{t,h} = \frac{1}{h} \sum_{i=0}^{h-1} H_{t+i}. \quad (1)$$

We also introduce the log return on the index $X_t = \log P_{t+1} - \log P_t$, in terms of which the gross yield can be rewritten

$$y_{t,h} = \frac{1}{h} \sum_{i=0}^{h-1} X_{t+i} + \frac{1}{h} \sum_{i=0}^{h-1} \log \left(1 + \frac{D_{t+i}}{P_{t+1+i}} \right), \quad (2)$$

where the first term on the right hand side readily reduces to $(\log P_{t+h} - \log P_t)/h$.

In the first part of our empirical work we regress yields on some explanatory variables that are known at the start of month t . We consider the following variables: the log earnings - price ratio $\log \langle e \rangle_t^{10} - \log P_t$, and the log earnings - price ratio centered on the prevailing long term interest rate $\log \langle e \rangle_t^{10} - \log P_t - \log r_t$. The quantity $\langle e \rangle_t^{10}$ refers to a moving average of real earnings over a time window of ten years. At variance with [Campbell and Shiller \(1988b\)](#), where the analysis is based on geometrical averages, we consider here arithmetical ones. The very reason is that in this way the current study can be readily adapted to the case of negative earnings. While an average over ten years resulting in a negative value is quite an unrealistic situation, it is not so uncommon to deal with spot realized earnings which are in fact losses. In this latter case the definition of $\log e_t$ is problematic. However, for the data set under

⁴Throughout this paper annual yields are obtained multiplying y_t by a factor 12.

consideration here this is never the case and we expect that moving from one average to the other does not change the overall picture. The use of an average of earnings in computing the ratios has been strongly pushed by the literature in recognition of the cyclical variability of earnings. In [Graham et al. \(1962\)](#) an approach is recommended that “shifts the original point of departure, or basis of computation, from the current earnings to the average earnings, which should cover a period of not less than five years, and preferably seven to ten years.”

In figures 1, 2, and 3 the solid lines represent regressions for the period 1871-2010, truncated where required in order to compute average over earnings and multi period returns. The solid points correspond to empirical data, while the role of dashed lines will be clarified later. In table 1 returns are measured over two until sixteen years (with the exclusion of odd years), as reported in the first column on a monthly basis. The second, third and fourth column give results of the regression of $(\log P_{t+h} - \log P_t)/h$ on $\log \langle e \rangle_t^{10} - \log P_t - \log r_t$ and $\log \langle e \rangle_t^{10} - \log P_t$, and of $y_{t,h}$ on $\log \langle e \rangle_t^{10} - \log P_t$, respectively; α and β correspond to the intercept and the precursor’s coefficient, while ϵ_α and ϵ_β are the associated standard errors. For each regression we report the R squared statistics and the significance level for an F test of the null that each coefficient is zero. Being the dependent variable a multi period return, we have corrected for autocorrelation effects employing the Newey-West correction, but we have not adjusted results for heteroscedasticity. All computational tasks have been performed with freely available software from the R project, <http://www.r-project.org/>. The table and figures show four relevant empirical facts: a) the striking ability of the log earning over price ratio to predict returns of the index, with an increasing R squared the longer the time horizon, b) this evidence increases when moving from returns to gross returns, c) when moving over different time horizons, the regression coefficients keep constant in a statistically robust way, and d) the pooriness of the prediction when the precursor is adjusted with long term interest rate. In light of the third point, we can ultimately explain the meaning of dashed lines in figures 1, 2, and 3. They correspond to linear curves whose coefficient are weighted averages of those obtained for fixed horizons, the weights being defined in terms of the R squared and of the standard error of each α_h and β_h (see section 4 for more details). Tables 2-4 report results for the same analysis performed over MSCI Indexes for eleven more countries. We can roughly group results in three

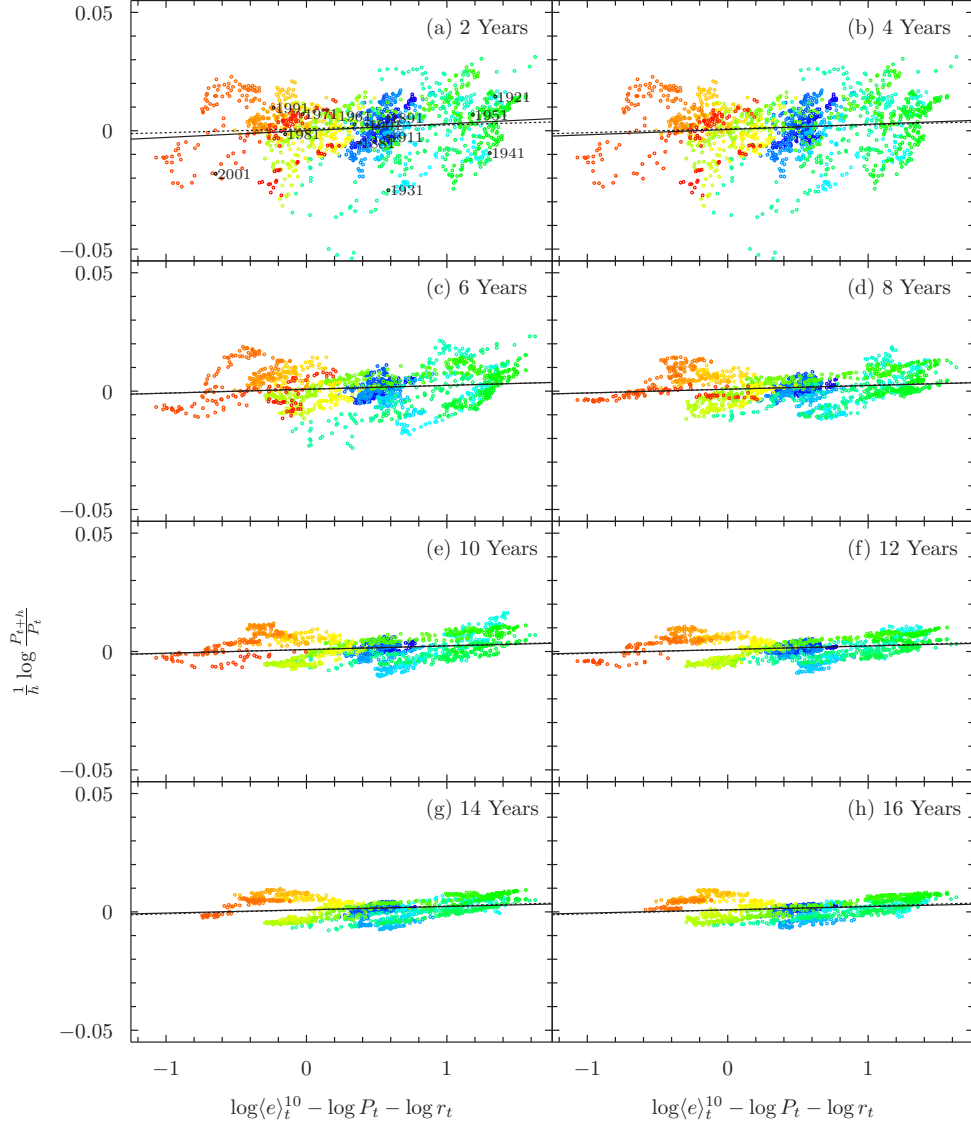


Figure 1: Regression of yield $(\log P_{t+h} - \log P_t)/h$ on the explanatory log earning - price ratio centered on r_t . Solid points: empirical data, solid line: linear regression, dashed line: averaged linear regression. Points are organized in chronological order according to the color scale ranging from dark blue to red passing through light blue, green, yellow, and orange; labels in the top left panel refer to points corresponding to the first month of the specified year.

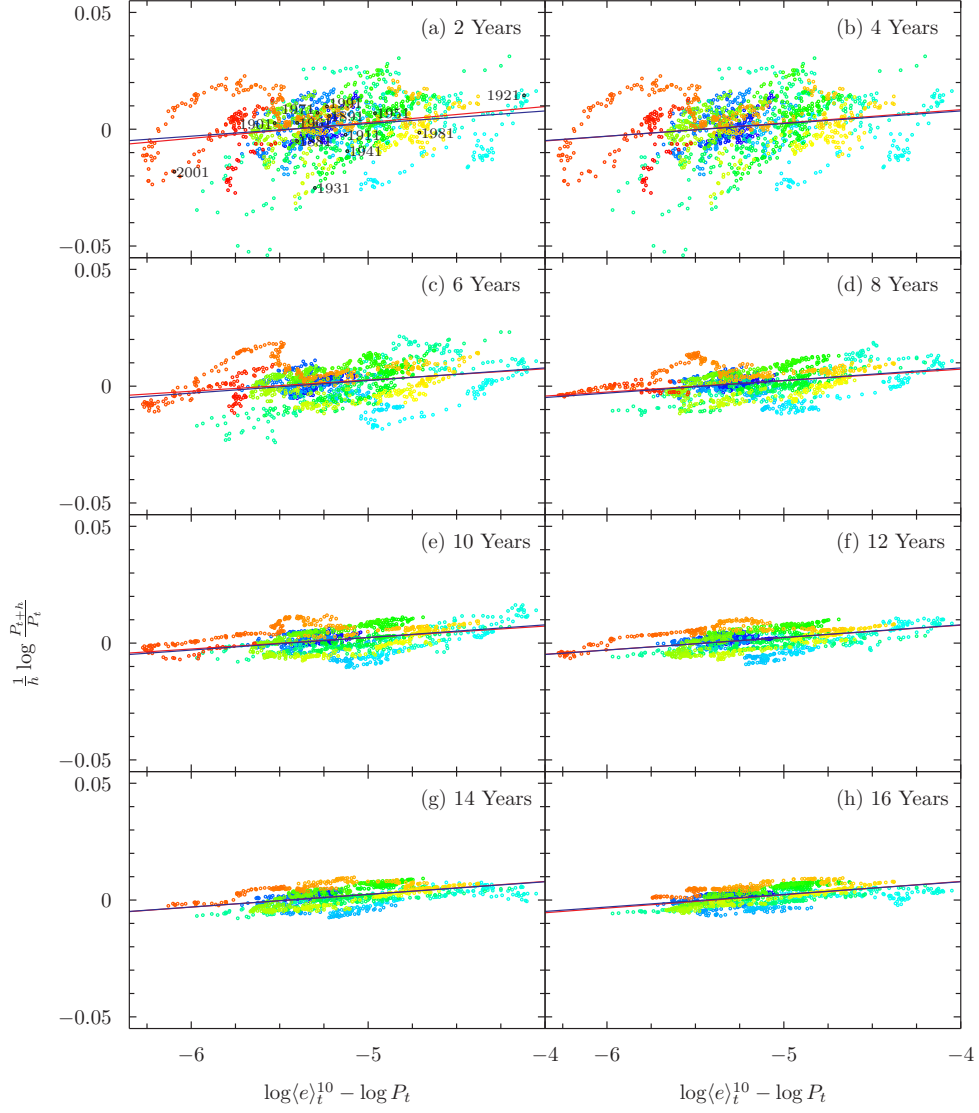


Figure 2: Regression of yield $(\log P_{t+h} - \log P_t)/h$ on the explanatory log earning - price ratio. Points, lines, labels, and color scale as in figure 1.

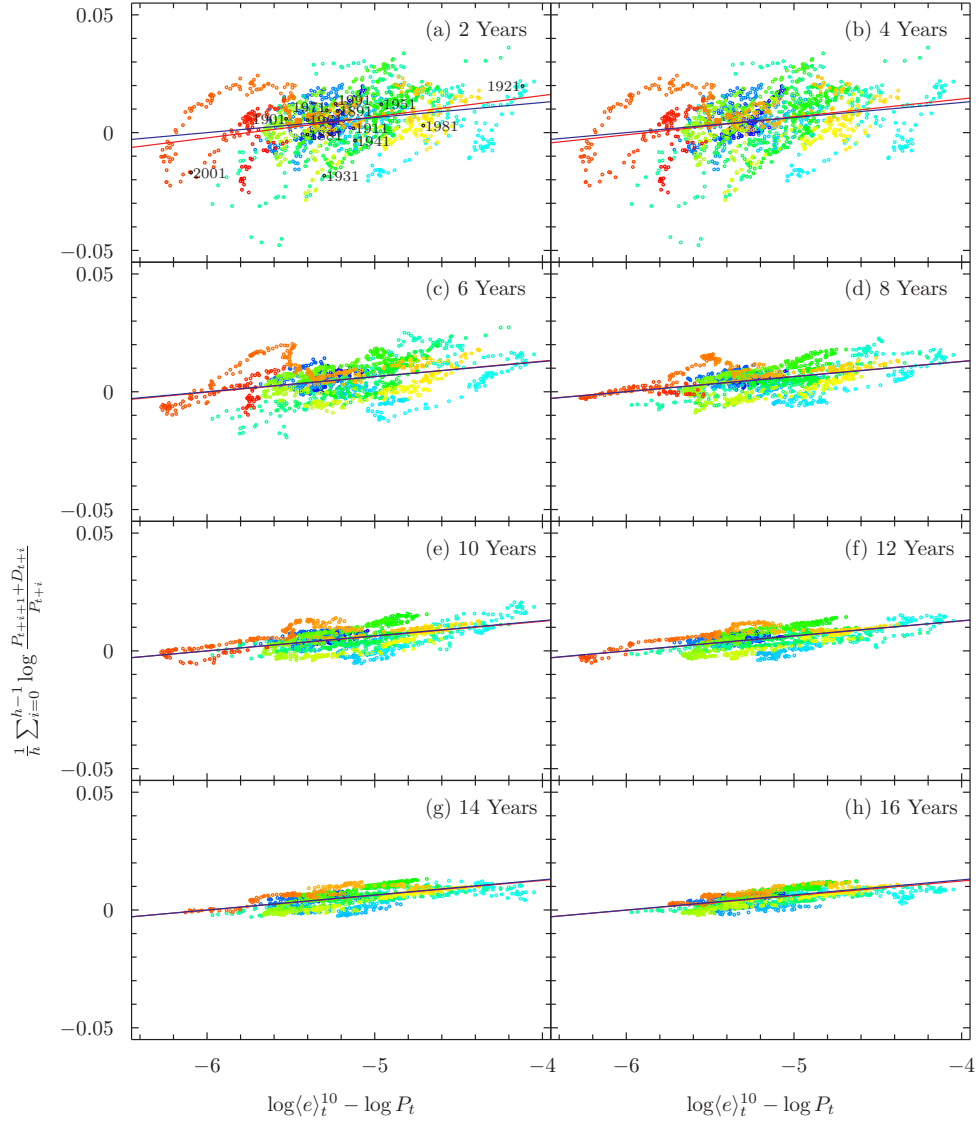


Figure 3: Regression of yield (1) on the explanatory log earning - price ratio. Points, lines, labels, and color scale as in figure 1.

Standard&Poor Composite, 1871-2010

h	$\frac{1}{h} \log \frac{P_{t+h}}{P_t} \text{ vs } \log \frac{\langle e \rangle_t^{10}}{P_t r_t}$			$\frac{1}{h} \log \frac{P_{t+h}}{P_t} \text{ vs } \log \frac{\langle e \rangle_t^{10}}{P_t}$			$y_{t,h} \text{ vs } \log \frac{\langle e \rangle_t^{10}}{P_t}$		
	$\alpha \pm \epsilon_\alpha$ Pr(>F)	$\beta \pm \epsilon_\beta$ Pr(>F)	R^2	$\alpha \pm \epsilon_\alpha$ Pr(>F)	$\beta \pm \epsilon_\beta$ Pr(>F)	R^2	$\alpha \pm \epsilon_\alpha$ Pr(>F)	$\beta \pm \epsilon_\beta$ Pr(>F)	R^2
24	0 ± 3 0.970	29 ± 5 0.198	2.1%	371 ± 36 0.023	69 ± 7 0.029	6.1%	518 ± 34 0.001	90 ± 7 0.003	10.8%
48	5 ± 2 0.676	22 ± 3 0.190	2.8%	313 ± 24 0.006	57 ± 6 0.009	9.5%	448 ± 22 —	76 ± 4 —	17.6%
72	8 ± 2 0.440	16 ± 3 0.227	2.7%	269 ± 18 0.001	49 ± 4 0.002	11.3%	394 ± 17 —	66 ± 3 —	22.0%
96	8 ± 2 0.383	16 ± 2 0.208	3.2%	269 ± 16 —	49 ± 3 —	15.2%	387 ± 14 —	65 ± 3 —	28.5%
120	8 ± 1 0.356	15 ± 2 0.181	3.8%	270 ± 14 —	49 ± 3 —	18.9%	379 ± 12 —	63 ± 2 —	33.9%
144	9 ± 1 0.328	15 ± 2 0.170	3.9%	289 ± 13 —	53 ± 3 —	23.4%	385 ± 11 —	65 ± 2 —	38.9%
168	9 ± 1 0.306	14 ± 2 0.186	4.0%	301 ± 12 —	55 ± 2 —	28.6%	379 ± 10 —	63 ± 2 —	43.5%
192	9 ± 1 0.305	14 ± 2 0.183	4.3%	308 ± 11 —	57 ± 2 —	35.3%	371 ± 9 —	62 ± 2 —	49.1%

Table 1: All values for α , β , and associated errors are expressed in 10^{-4} units; p-values of order or smaller than 10^{-4} have been left blank.

main categories: impressive forecasting power with a fraction of explained variance reaching levels of order 90% (Belgium, France, Germany, Japan, the Netherlands, United Kingdom); very poor forecasting capability and modest or negligible R squared level (Australia and Canada); countries (Norway, Sweden, Switzerland) with an overall good performance, but for which the fraction of explained variance behaves anomalously for long horizons (see ten and twelve years returns for Norway, and twelve years returns for Sweden and Switzerland). The increase of R squared induced by gross returns is also confirmed, even though this evidence is less strong than for the case of US data. It is worth to stress that the difference of sample size between US and remaining countries data surely plays a role which is however hard to quantify. It is far from the aim of this work to provide an economic explanation of the behaviour of Australia and Canada. In the next section we will introduce a dynamical model able to reproduce the

MSCI Australia, 1970-2010				MSCI Belgium, 1970-2010					
h	Log returns		Gross log returns		Log returns		Gross log returns		
	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2
24	$-93 \pm 92^{**}$	$-24 \pm 18^{**}$	0.6%	$-44 \pm 91^{**}$	$-21 \pm 18^{**}$	0.4%	$729 \pm 97^{**}$	$124 \pm 18^{**}$	12.7%
48	$111 \pm 76^{**}$	$16 \pm 15^{**}$	0.4%	$168 \pm 76^{**}$	$21 \pm 15^{**}$	0.7%	942 ± 54	165 ± 10	46.5%
72	$220 \pm 55^*$	$39 \pm 11^{**}$	4.5%	$287 \pm 54^*$	$46 \pm 11^*$	6.2%	808 ± 24	140 ± 5	75.0%
96	$44 \pm 33^{**}$	$5 \pm 7^{**}$	0.2%	$113 \pm 33^{**}$	$12 \pm 7^{**}$	1.4%	580 ± 19	98 ± 4	73.9%
120	$28 \pm 28^{**}$	$2 \pm 6^{**}$	0.1%	$93 \pm 28^{**}$	$9 \pm 6^{**}$	1.1%	493 ± 16	82 ± 3	75.8%
144	$38 \pm 29^{**}$	$4 \pm 6^{**}$	0.3%	$96 \pm 28^{**}$	$10 \pm 6^{**}$	1.4%	449 ± 13	74 ± 2	81.5%
168	$31 \pm 29^{**}$	$2 \pm 6^{**}$	0.1%	$83 \pm 28^{**}$	$6 \pm 6^{**}$	0.1%	426 ± 9	70 ± 2	90.9%
192	$-90 \pm 22^*$	$-23 \pm 5^{**}$	14.7%	$-46 \pm 21^{**}$	$-21 \pm 4^*$	13.4%	318 ± 8	50 ± 2	86.5%

MSCI Canada, 1970-2010				MSCI France, 1970-2010					
h	Log returns		Gross log returns		Log returns		Gross log returns		
	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2
24	$136 \pm 58^{**}$	$17 \pm 10^{**}$	0.8%	$229 \pm 56^{**}$	$30 \pm 10^{**}$	2.6%	$495 \pm 65^{**}$	$75 \pm 12^{**}$	11.4%
48	$70 \pm 39^{**}$	$5 \pm 7^{**}$	0.2%	$154 \pm 38^{**}$	$17 \pm 7^{**}$	2.0%	618 ± 29	98 ± 5	53.5%
72	$22 \pm 27^{**}$	$-3 \pm 5^*$	0.1%	$100 \pm 26^*$	$7 \pm 5^{**}$	0.9%	487 ± 17	76 ± 3	68.0%
96	$-17 \pm 22^{**}$	$-10 \pm 4^{**}$	2.6%	$59 \pm 20^{**}$	$0 \pm 4^{**}$	0.0%	404 ± 10	62 ± 2	81.4%
120	-126 ± 19	-32 ± 4	26.3%	$-42 \pm 17^{**}$	-20 ± 3	14.4%	343 ± 9	52 ± 2	80.9%
144	-166 ± 20	-41 ± 4	37.1%	$-79 \pm 18^*$	-28 ± 3	24.9%	301 ± 8	44 ± 2	80.6%
168	$-84 \pm 24^{**}$	$-25 \pm 5^*$	14.6%	$-1 \pm 22^{**}$	$-13 \pm 4^{**}$	5.2%	275 ± 8	40 ± 2	78.7%
192	$79 \pm 26^*$	$7 \pm 5^{**}$	1.3%	149 ± 23	$17 \pm 5^{**}$	8.5%	248 ± 7	35 ± 1	80.8%

MSCI Germany, 1970-2010					MSCI Japan, 1970-2010				
h	Log returns		Gross log returns		Log returns		Gross log returns		
	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2
24	978 \pm 98*	161 \pm 18*	21.3%	1050 \pm 96*	170 \pm 17*	23.7%	827 \pm 113*	127 \pm 18*	13.9%
48	976 \pm 48	162 \pm 9	54.6%	1038 \pm 47	169 \pm 8	57.8%	978 \pm 81	152 \pm 13	32.6%
72	825 \pm 26	137 \pm 5	76.4%	884 \pm 26	143 \pm 5	77.9%	1005 \pm 56	159 \pm 9	54.6%
96	573 \pm 19	92 \pm 3	74.1%	633 \pm 20	99 \pm 4	75.0%	828 \pm 37	133 \pm 6	67.0%
120	474 \pm 17	75 \pm 3	73.4%	532 \pm 17	82 \pm 3	74.9%	567 \pm 27	92 \pm 4	67.2%
144	448 \pm 14	71 \pm 3	79.3%	502 \pm 14	77 \pm 3	81.7%	526 \pm 20	86 \pm 3	78.0%
168	439 \pm 12	70 \pm 2	86.0%	486 \pm 11	75 \pm 2	88.2%	515 \pm 12	84 \pm 2	91.9%
192	371 \pm 11	58 \pm 2	85.0%	412 \pm 10	61 \pm 2	87.3%	394 \pm 11	64 \pm 2	89.2%

MSCI Netherlands, 1970-2010					MSCI Norway, 1970-2010				
h	Log returns		Gross log returns		Log returns		Gross log returns		
	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2
24	710 \pm 69*	120 \pm 13*	20.7%	815 \pm 68	134 \pm 13*	24.9%	1339 \pm 110	245 \pm 21*	30.5%
48	794 \pm 39	137 \pm 8	52.7%	887 \pm 39	149 \pm 8	57.1%	803 \pm 62	144 \pm 12	33.9%
72	717 \pm 22	123 \pm 4	76.4%	805 \pm 22	135 \pm 4	78.3%	460 \pm 41	80 \pm 8	28.3%
96	559 \pm 16	94 \pm 3	79.9%	643 \pm 17	105 \pm 3	81.3%	354 \pm 31	60 \pm 6	29.7%
120	485 \pm 13	81 \pm 2	82.8%	565 \pm 13	90 \pm 3	84.2%	92 \pm 20**	10 \pm 4**	3.0%
144	454 \pm 11	75 \pm 2	86.0%	528 \pm 11	84 \pm 2	88.1%	174 \pm 32*	26 \pm 6**	8.2%
168	442 \pm 8	74 \pm 2	91.4%	508 \pm 8	81 \pm 2	93.4%	348 \pm 32	61 \pm 6	35.0%
192	372 \pm 8	61 \pm 2	89.4%	428 \pm 8	66 \pm 2	90.7%	278 \pm 28	47 \pm 5	33.6%

Table 3: All values for α , β , and associated errors are expressed in 10^{-4} units. Significance codes: 0 ‘’, 0.01 ‘*’, 0.1 ‘**’.

empirical evidences shared by the majority of countries we have just analyzed.

3 The model

In the following we will prove some results concerning a dynamics driven both by a value investment and a momentum component. In the first instance we will consider only the role played by stock prices, but later we will take into account gross returns and the effect of dividends. Our model rests on the following two assumptions inspired by the evidences identified in the previous section.

Assumption 1 *Financial returns X_t are naturally driven by the growth rate μ_t , and by an ancillary process ξ_t satisfying*

$$\xi_{t+1} = \xi_t + \frac{\kappa}{1-\gamma} \sigma_X W_t^X, \quad (3)$$

with initial time condition $\xi_0 = 0$. The ξ_t process ensures the diffusive behaviour of stock prices. As will be clear later, the positive parameter σ_X corresponds to the long run returns volatility, while the prefactor depending on $\kappa > 0$ and $0 < \gamma < 1$ has been added for future convenience. The W_t^X 's for $t = 0, \dots, h$ are independent identically distributed (i.i.d.) standard Gaussian increments.

Assumption 2 *The dynamics of log returns' growth rate is given in terms of the difference equation*

$$\mu_{t+1} = \gamma \mu_t + \kappa [\log \langle e \rangle_t^{10} - X_t + \mathcal{H} + g \mathcal{F} (\log \langle e \rangle_0^{10} - \log P_0) t] + \sigma_\mu W_t^\mu \quad (4)$$

where the parameter σ_μ is a positive volatility constant, and the W_t^μ 's for $t = 0, \dots, h$ are i.i.d. standard Gaussian increments, not dependent on the W_t^X 's. The return growth μ_t depends on its own value in the previous period, with γ agents' sensitivity to market trend. This momentum effect can be naturally justified in terms of agents' expectation that returns are higher in bullish markets than in bearish ones. The "fundamental" component is proportional to a function \mathcal{F} of the level of the logarithmic averaged earnings over price ratio at time zero, where the proportionality constant is given by the rate of earnings growth g , implicitly defined by $\langle e \rangle_t^{10} =$

MSCI Sweden, 1970-2010				MSCI Switzerland, 1970-2010			
h	Log returns			Log returns			Gross log returns
	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$ $\beta \pm \epsilon_\beta$ R^2
24	1224 \pm 97	71 \pm 12*	9.9%	453 \pm 67*	71 \pm 12*	9.9%	510 \pm 66 78 \pm 12* 12.1%
48	824 \pm 56	74 \pm 8	23.1%	471 \pm 44	74 \pm 8	23.1%	521 \pm 43 80 \pm 8 27.1%
72	657 \pm 36	55 \pm 6	25.1%	364 \pm 32	55 \pm 6	25.1%	408 \pm 31 60 \pm 6 30.0%
96	570 \pm 26	39 \pm 5	22.7%	274 \pm 25	39 \pm 5	22.7%	315 \pm 24 43 \pm 4 28.4%
120	275 \pm 24	24 \pm 4**	13.0%	196 \pm 22*	24 \pm 4**	13.0%	238 \pm 21 29 \pm 4* 19.1%
144	206 \pm 23	5 \pm 3**	1.2%	100 \pm 17*	5 \pm 3**	1.2%	148 \pm 17 11 \pm 3** 6.0%
168	294 \pm 23	17 \pm 3**	12.7%	160 \pm 17	17 \pm 3**	12.7%	206 \pm 17 22 \pm 3* 21.3%
192	379 \pm 20	47 \pm 3	65.5%	315 \pm 15	47 \pm 3	65.5%	354 \pm 14 52 \pm 3 71.6%

MSCI United Kingdom, 1970-2010				MSCI United Kingdom, 1970-2010			
h	Log returns			Log returns			Gross log returns
	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$ $\beta \pm \epsilon_\beta$ R^2
24	382 \pm 76**	67 \pm 15**	6.1%	464 \pm 75**	77 \pm 15**	8.1%	464 \pm 75** 77 \pm 15** 8.1%
48	514 \pm 45	92 \pm 9	27.5%	589 \pm 44	101 \pm 9	32.0%	589 \pm 44 101 \pm 9 32.0%
72	422 \pm 22	75 \pm 4	52.8%	495 \pm 22	83 \pm 4	58.8%	495 \pm 22 83 \pm 4 58.8%
96	279 \pm 12	48 \pm 2	61.4%	354 \pm 12	56 \pm 2	68.0%	354 \pm 12 56 \pm 2 68.0%
120	219 \pm 10	36 \pm 2	57.7%	294 \pm 11	45 \pm 2	66.6%	294 \pm 11 45 \pm 2 66.6%
144	198 \pm 14	32 \pm 3	39.0%	270 \pm 14	40 \pm 3	51.9%	270 \pm 14 40 \pm 3 51.9%
168	234 \pm 14	40 \pm 3	53.6%	299 \pm 13	47 \pm 3	65.0%	299 \pm 13 47 \pm 3 65.0%
192	214 \pm 9	36 \pm 2	71.9%	270 \pm 8	41 \pm 2	78.9%	270 \pm 8 41 \pm 2 78.9%

Table 4: All values for α , β , and associated errors are expressed in 10^{-4} units. Significance codes: 0 ‘’, 0.01 ‘*’, 0.1 ‘**’.

$\langle e \rangle_0^{10} \exp(gt)$. This component grows linearly with time, but we also allow for a correction equal to \mathcal{H} , possibly function of the initial time log EP ratio. The initial time value of log EP determines the reference growth level, from which the actual stock price may deviate as an effect of random external disturbances. Investors reallocate assets in response to this disequilibrium causing stock prices to move in the direction that reduces the deviation. The full adjustment is not immediate but it takes a typical time κ^{-1} the return to mean revert to the fundamental level.

In conclusion, our model states the dynamics of log returns can be fully specified by the linear stochastic difference system

$$\begin{cases} X_{t+1} = X_t + \mu_t + \xi_t \\ \mu_{t+1} = \gamma\mu_t + \kappa(\log\langle e \rangle_t^{10} - X_t + \mathcal{H} + g\mathcal{F}t) + \sigma_\mu W_t^\mu \\ \xi_{t+1} = \xi_t + \frac{\kappa}{1-\gamma}\sigma_X W_t^X \end{cases}, \quad (5)$$

with $X_0 = \log P_0$ and μ_0 initial time conditions. By a further change of variable $Y_t = X_t - \log\langle e \rangle_0^{10}$, the expected values $\mathbb{E}_0[Y_t]$, $\mathbb{E}_0[\mu_t]$, $\mathbb{E}_0[\xi_t]$ conditional at the information available at time 0 evolve according to the linear first order difference system

$$\begin{cases} \mathbb{E}_0[Y_{t+1}] = \mathbb{E}_0[Y_t] + \mathbb{E}_0[\mu_t] + \mathbb{E}_0[\xi_t] \\ \mathbb{E}_0[\mu_{t+1}] = -\kappa\mathbb{E}_0[Y_t] + \gamma\mathbb{E}_0[\mu_t] + \kappa\mathcal{H} + \kappa g(1 + \mathcal{F})t \\ \mathbb{E}_0[\xi_{t+1}] = \mathbb{E}_0[\xi_t] \end{cases},$$

with the initial conditions

$$\begin{cases} \mathbb{E}_0[Y_0] = X_0 - \log\langle e \rangle_0^{10} \\ \mathbb{E}_0[\mu_0] = \mu_0 \\ \mathbb{E}_0[\xi_0] = 0 \end{cases}.$$

Lemma 1. *The expected rate of return is linear in $\mathcal{F}(\log\langle e \rangle_0^{10} - \log P_0)$, provided we consider a sufficiently large number of periods h*

$$\mathbb{E}_0 \left[\frac{1}{h} \log \frac{P_{0+h}}{P_0} \right] = g(1 + \mathcal{F}) + O\left(\frac{1}{h}\right). \quad (6)$$

Moreover its variance converges to zero with rate of convergence equal to minus one

$$\text{Var}_0 \left[\frac{1}{h} \log \frac{P_{0+h}}{P_0} \right] = \frac{\sigma_X^2}{h} + o\left(\frac{1}{h}\right). \quad (7)$$

Proof. The stochastic dynamical system (5) may be put in the vector form

$$\mathbf{V}_{t+1} = \mathbb{J} \mathbf{V}_t + \kappa \mathcal{H} \mathbf{e}_2 + \kappa g(1 + \mathcal{F})t \mathbf{e}_2 + \sigma_\mu W_t^\mu \mathbf{e}_2 + \frac{\kappa}{1 - \gamma} \sigma_X W_t^X \mathbf{e}_3 \quad (8)$$

where $\mathbb{J} = \begin{bmatrix} 1 & 1 & 1 \\ -\kappa & \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is the matrix of the systems, \mathbf{e}_i for $i = 1, 2, 3$ is the orthonormal canonical base, and $\mathbf{V}_t = \begin{bmatrix} Y_t \\ \mu_t \\ \xi_t \end{bmatrix}$. The set of the eigenvalues of \mathbb{J} includes 1, independently of κ and γ , and two ones more

$$\begin{aligned} \lambda_+ &= \frac{\gamma + 1}{2} + \frac{1}{2} \sqrt{(1 - \gamma)^2 - 4\kappa}, \\ \lambda_- &= \frac{\gamma + 1}{2} - \frac{1}{2} \sqrt{(1 - \gamma)^2 - 4\kappa}. \end{aligned}$$

In order to ensure $|\lambda_+| < 1$ and $|\lambda_-| < 1$, we need to impose the following restrictions

$$0 < \gamma < 1, \quad \text{and} \quad 0 < \kappa \leq \frac{(1 - \gamma)^2}{4}. \quad (9)$$

Iterating (8) we obtain

$$\begin{aligned} \mathbf{V}_h &= \mathbb{J}^h \mathbf{V}_0 + \kappa [g(1 + \mathcal{F})(h - 1) + \mathcal{H}] \sum_{k=0}^{h-1} \mathbb{J}^k \mathbf{e}_2 - \kappa g(1 + \mathcal{F}) \sum_{k=0}^{h-1} k \mathbb{J}^k \mathbf{e}_2 \\ &\quad + \sigma_\mu \sum_{k=0}^{h-1} W_k^\mu \mathbb{J}^{h-1-k} \mathbf{e}_2 + \frac{\kappa}{1 - \gamma} \sigma_X \sum_{k=0}^{h-1} W_k^X \mathbb{J}^{h-1-k} \mathbf{e}_3. \end{aligned}$$

We can rewrite more conveniently $\mathbb{J} = \mathbb{Q} \Lambda \mathbb{Q}^{-1}$ by means of the matrices $\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda_+ & 0 \\ 0 & 0 & \lambda_- \end{bmatrix}$,

$$\mathbb{Q} = \begin{bmatrix} 1-\gamma & 1 & 1 \\ -\kappa & \lambda_+ - 1 & \lambda_- - 1 \\ \kappa & 0 & 0 \end{bmatrix}, \text{ and } \mathbb{Q}^{-1} = \frac{1}{|\mathbb{Q}|} \begin{bmatrix} 0 & 0 & \lambda_- - \lambda_+ \\ \kappa(\lambda_- - 1) & -\kappa & (1-\gamma)(1-\lambda_-) - \kappa \\ \kappa(1-\lambda_+) & \kappa & \kappa - (1-\gamma)(1-\lambda_+) \end{bmatrix}$$

with $|\mathbb{Q}| = \kappa(\lambda_- - \lambda_+)$. Taking expectation of $\mathbf{e}_1^t \mathbf{V}_h$ we obtain

$$\begin{aligned} \mathbb{E}_0[Y_h] &= \mathbf{e}_1^t \mathbb{Q} \Lambda^h \mathbb{Q}^{-1} \mathbf{V}_0 \\ &\quad + \kappa [g(1+\mathcal{F})(h-1) + \mathcal{H}] \sum_{k=0}^{h-1} \mathbf{e}_1^t \mathbb{Q} \Lambda^k \mathbb{Q}^{-1} \mathbf{e}_2 - \kappa g(1+\mathcal{F}) \sum_{k=0}^{h-1} k \mathbf{e}_1^t \mathbb{Q} \Lambda^k \mathbb{Q}^{-1} \mathbf{e}_2 \\ &= -\frac{\lambda_+^h}{\lambda_- - \lambda_+} [(1-\lambda_-)Y_0 + \mu_0] + \frac{\lambda_-^h}{\lambda_- - \lambda_+} [(1-\lambda_+)Y_0 + \mu_0] \\ &\quad + g(1+\mathcal{F})(h-1) + \mathcal{H} - [g(1+\mathcal{F}) - \mathcal{H}] \frac{\kappa}{\lambda_- - \lambda_+} \left(\frac{\lambda_+^h}{1-\lambda_+} - \frac{\lambda_-^h}{1-\lambda_-} \right) \\ &\quad - g(1+\mathcal{F}) \frac{\kappa}{\lambda_- - \lambda_+} \left[\lambda_- \frac{1-\lambda_-^h}{(1-\lambda_-)^2} - \lambda_+ \frac{1-\lambda_+^h}{(1-\lambda_+)^2} \right]. \end{aligned} \quad (10)$$

From previous expression relation (6) immediately follows.

As far as the variance of Y_h is concerned, we have

$$\text{Var}_0[Y_h] = \sum_{k=0}^{h-1} \left[\sigma_\mu^2 (\mathbf{e}_1^t \mathbb{Q} \Lambda^k \mathbb{Q}^{-1} \mathbf{e}_2)^2 + \frac{\kappa^2}{(1-\gamma)^2} \sigma_X^2 (\mathbf{e}_1^t \mathbb{Q} \Lambda^k \mathbb{Q}^{-1} \mathbf{e}_3)^2 \right], \quad (11)$$

from which we find

$$\begin{aligned} \text{Var}_0[Y_h] &= \frac{\sigma_\mu^2}{(\lambda_- - \lambda_+)^2} \sum_{k=0}^{h-1} [\lambda_-^k - \lambda_+^k]^2 \\ &\quad + \frac{\sigma_X^2}{(\lambda_- - \lambda_+)^2} \sum_{k=0}^{h-1} \left[\lambda_- - \lambda_+ + \lambda_+^k (1 - \lambda_- - \frac{\kappa}{1-\gamma}) + \lambda_-^k (\frac{\kappa}{1-\gamma} - 1 + \lambda_+) \right]^2, \end{aligned}$$

and the thesis (7) follows. \square

It is a posteriori clear why we have fixed the proportionality constant of the fundamental component equal to g . Indeed, in case g was zero, not only averaged earnings but also stock returns would grow sub-exponentially, which is quite reasonable. The variance of Y_h increases linearly with h , which is a diffusive behavior of stock returns any realistic model should satisfy at the leading order. Finer effects, like possible return non zero autocorrelations induced by business cycles over long horizons could be included in our model modifying the covariance structure of the $\{W_t^X\}$ noise. Moreover it is worth to remark that none of our result crucially rest on the Gaussian nature of noise increments.

From previous Lemma, the interesting result we state below follows.

Corollary 2. *If \mathcal{F} is linear in $\log\langle e \rangle_0^{10} - \log P_0$, then the model (5) is able to reproduce the linear scaling of returns with the initial log EP ratio on the long run.*

The next step forward is to add a dividend components to our stock returns model. In this respect we mainly follow the proposal discussed in [Campbell and Shiller \(1988a,b\)](#), where it is argued that the log dividend price ratio follows a stationary stochastic process, and the fixed mean of $\log D_t - \log P_t$, $\log \mathcal{G}$, can be used as an expansion point

$$\Delta(d_{t-1} - X_t) = -\theta(d_{t-1} - X_t - \log \mathcal{G}) + \sigma_d W_t^d, \quad (12)$$

with $d_t = \log D_t$, $\sigma_d > 0$, and $\{W_t^d\}$ for $t = 1, \dots, h$ i.i.d. standard Gaussian variates. At variance with the proposal of Campbell and Shiller, we also assume that \mathcal{G} can possibly depend on the log EP at time zero, i.e. $\mathcal{G} = \mathcal{G}(\log\langle e \rangle_0^{10} - \log P_0)$.

Lemma 3. *In the limit $h \gg 1$ the quantity $\frac{1}{h} \sum_{i=0}^{h-1} \log \left(1 + \frac{D_i}{P_{i+1}}\right)$ contributes to the rate of log gross returns proportionally to \mathcal{G} , and its variance converges to zero with rate minus one.*

Proof. On a monthly basis the quantity D_t/P_{t+1} is of order 10^{-3} , and we are allowed to replace $\log \left(1 + \frac{D_i}{P_{i+1}}\right)$ with D_i/P_{i+1} . In the same spirit of [Campbell and Shiller \(1988a,b\)](#), in order to prove the thesis we need to Taylor expand quantities of interest around $\log \mathcal{G}$. In particular, we have

$$\frac{D_i}{P_{i+1}} = \exp(d_i - X_{i+1}) \simeq \exp(\log \mathcal{G}) (1 + d_i - X_{i+1} - \log \mathcal{G}).$$

Iterating equation (12) and taking expectation, we obtain for $i \geq 0$

$$\mathbb{E}_0[d_{i-1} - X_i] = (1 - \theta)^i(d_{-1} - X_0) + \log \mathcal{G} [1 - (1 - \theta)^i], \quad (13)$$

where d_{-1} is the log dividend prevailing at time zero being cumulated between time -1 and zero. Under the constraint $0 < \theta < 2$, we can conclude

$$\begin{aligned} \frac{1}{h} \mathbb{E}_0 \left[\sum_{i=0}^{h-1} \log \left(1 + \frac{D_i}{P_{i+1}}\right) \right] &\simeq \mathcal{G}(1 - \log \mathcal{G}) + \mathcal{G} \log \mathcal{G} + \mathcal{G} \frac{1 - \theta}{\theta} (\log DP_0 - \log \mathcal{G}) \frac{1 - (1 - \theta)^h}{h} \\ &= \mathcal{G} + O\left(\frac{1}{h}\right), \end{aligned}$$

and

$$\text{Var}_0 \left[\frac{1}{h} \sum_{i=0}^{h-1} \log \left(1 + \frac{D_i}{P_{i+1}} \right) \right] \simeq \frac{1}{h} \frac{\mathcal{G}^2 \sigma_d^2}{\theta(2-\theta)} + o\left(\frac{1}{h}\right).$$

□

The expansion performed in the above proof is slightly different from that proposed in [Campbell and Shiller \(1988a,b\)](#), where $\log(P_t + D_t)$ is approximated by $\vartheta \log P_t + (1 - \vartheta) \log D_t + \varkappa$ with $\vartheta = 1/(1 + \mathcal{G})$ and $\varkappa = \log(1 + \mathcal{G}) - \vartheta \mathcal{G} \log \mathcal{G}$. In the same spirit of the fourth point in the Appendix of [Campbell and Shiller \(1988a\)](#), in the next section we will evaluate the impact of the approximation on the quality of regressions.

Corollary 4. *If \mathcal{G} is linear in $\log \langle e \rangle_0^{10} - \log P_0$, then equation (12) is able to reproduce the linear scaling of dividends' contribution to the yield with the initial log EP ratio on the long run.*

We are now ready to state the main theoretical result of this paper.

Proposition 5. *The expected gross yield (1) is linear in \mathcal{F} and \mathcal{G} , provided we consider sufficiently long horizons*

$$\mathbb{E}_0[y_h] \simeq g(1 + \mathcal{F}) + \mathcal{G} + O\left(\frac{1}{h}\right).$$

The proof is an immediate consequence of the Corollaries 2 and 4. More interestingly, from the proofs of previous Lemmas we can explicitly identify the contribution to the long term yield whose scaling over time is more persistent, and which ultimately determines its convergence towards the limit $g(1 + \mathcal{F}) + \mathcal{G}$. The expression of the leading correction proportional to one over h reads

$$\mathcal{H} - g(1 + \mathcal{F}) \left[1 + \kappa \frac{1 - \lambda_- \lambda_+}{(1 - \lambda_-)^2 (1 - \lambda_+)^2} \right] + \log EP_0 + \mathcal{G} \left(1 - \frac{1}{\theta} \right) (\log DP_0 - \log \mathcal{G}). \quad (14)$$

From equation (13) we know that $(1 - \theta)^i$ plays the role of a damping function. Assuming $\theta \ll 1$, it reduces to $\exp(-i/\tau_\theta)$ where the quantity $\tau_\theta = 1/\theta$ plays the role of the typical time scale of the mean-reverting process. We reasonably expect that the last term in the expression (14) does not contribute too much to the leading correction, since we neither expect

an extremely large time scale for the process nor an extreme discrepancy between $\log DP_0$ and $\log \mathcal{G}$. Another interesting point to note is the choice of μ_0 plays a minor role in relation to the speed of convergence of the process to the long term expected value. Indeed, μ_0 does not appear in expression (14), while from (10) we know that its effect is exponentially damped by λ_-^h and λ_+^h . Being $0 < \lambda_- < \lambda_+ < 1$, the dominating contribution for $h \gg 1$ is given by λ_+^h , implying a typical damping scale equal to $-(\log \lambda_+)^{-1}$.

4 Parameters calibration

Corollaries 2 and 4 state that the linear scaling of long term yield with the level of log EP ratio can be reproduced by our dynamical model if we assume for the quantities $g(1 + \mathcal{F})$ and \mathcal{G} the form $g(1 + \mathcal{F}) = \alpha_{\mathcal{F}} + \beta_{\mathcal{F}} \log EP_0$ and $\mathcal{G} = \alpha_{\mathcal{G}} + \beta_{\mathcal{G}} \log EP_0$. This suggest an immediate and easy way to fix the values of $\alpha_{\mathcal{F}}$, $\beta_{\mathcal{F}}$, $\alpha_{\mathcal{G}}$, and $\beta_{\mathcal{G}}$. In Tables 2 - 4 we have reported the values of α_h , ϵ_{α_h} , β_h , and ϵ_{β_h} , with the index $h = 24 \times i$ for $i = 1, \dots, 8$ corresponding to the number of elapsed months, obtained from the regression of $(\log P_{t+h} - \log P_t) / h$ on $\log \langle e \rangle_t^{10} - \log P_t$. The values for $\alpha_{\mathcal{F}}$ and $\epsilon_{\alpha_{\mathcal{F}}}$ are then computed by

$$\alpha_{\mathcal{F}} = \sum_h w_h \alpha_h \quad \text{with} \quad w_h = \frac{|\alpha_h|}{\epsilon_{\alpha_h}} \frac{R_h^2}{\sum_h \frac{|\alpha_h|}{\epsilon_{\alpha_h}} R_h^2}, \quad \text{and} \quad \epsilon_{\alpha_{\mathcal{F}}} = \sum_h w_h \epsilon_{\alpha_h},$$

and similarly for $\beta_{\mathcal{F}}$ and the associated error. The weights have been chosen in order not only to privilege regression's coefficients with the lowest relative error $\epsilon_{\alpha_h}/|\alpha_h|$, but also those with the highest R squared. The quantities labelled with \mathcal{G} are obtained in the same way after regressing the dividends' contribution to the yield $\frac{1}{h} \sum_{i=0}^{h-1} \log \left(1 + \frac{D_{t+i}}{P_{t+1+i}} \right)$ on the initial log EP. Equipped with the values of $\alpha_{\mathcal{G}}$, and $\beta_{\mathcal{G}}$, see Table 5, in Table 6 we report the comparison between the regression of $\frac{1}{h} \mathbb{E}_0 \left[\sum_{i=0}^{h-1} \log \left(1 + \frac{D_i}{P_{i+1}\mathcal{G}} \right) \right]$ and the approximate expression $\mathcal{G} \frac{1}{h} \mathbb{E}_0 \left[\sum_{i=0}^{h-1} \left(1 + \log \frac{D_i}{P_{i+1}\mathcal{G}} \right) \right]$ on information provided by the log EP ratio. We explicitly detail cases of Belgium, France, United Kingdom and Germany, but all other countries perform exactly in the same way. Results are in full statistical agreement, with a strong level of significance. There are small differences, but they do not invalidate the overall picture. Once

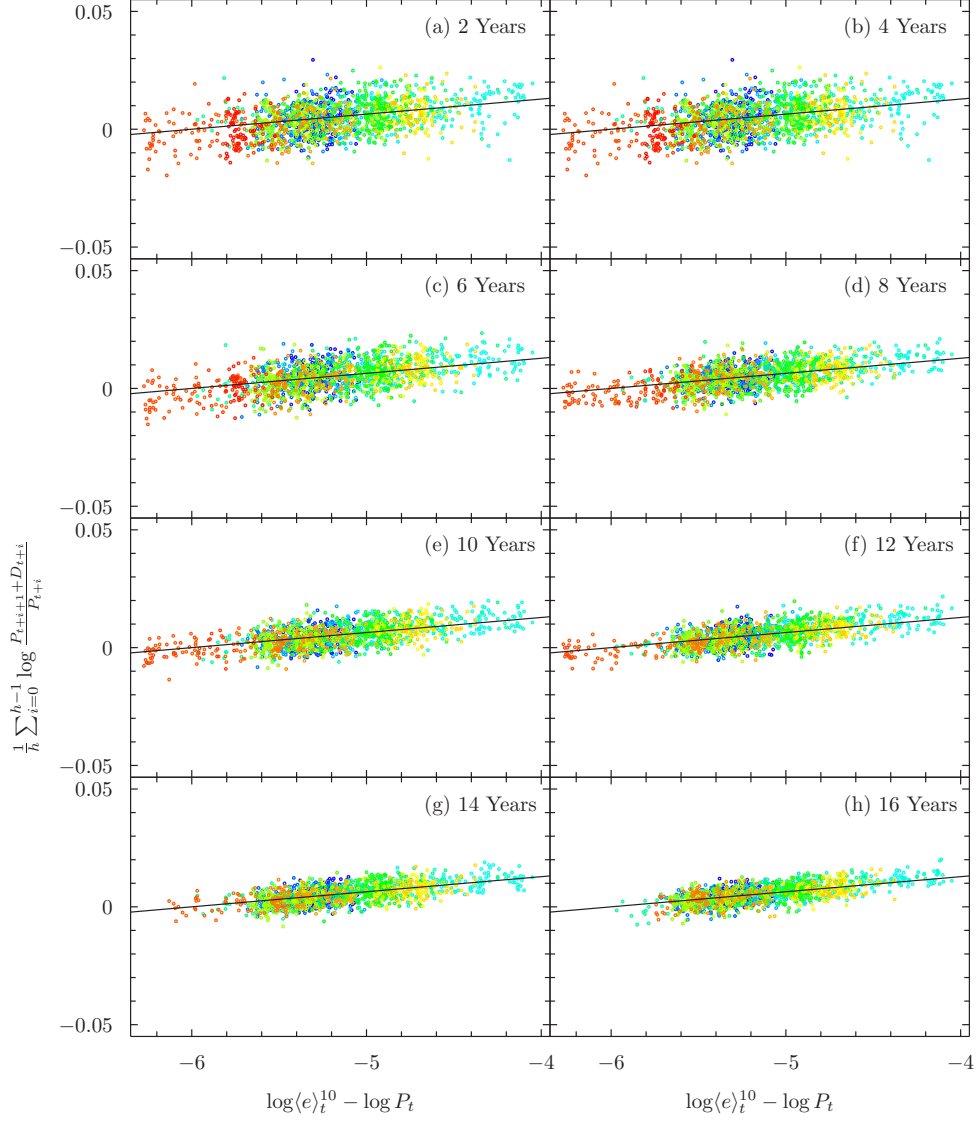


Figure 4: Monte Carlo scenarios generated simulating the model (5)-(12) with initial time conditions Y_0 , μ_0 , and $\log DP_0$ equal to the empirical ones; the solid line corresponds to the linear regression with coefficient as in Table 5. Color scale as in Figures 1-3.

	$g(1 + \mathcal{F}) = \alpha_{\mathcal{F}} + \beta_{\mathcal{F}} \log \text{EP}_0$		$\mathcal{G} = \alpha_{\mathcal{G}} + \beta_{\mathcal{G}} \log \text{EP}_0$	
	$\alpha_{\mathcal{F}}$	$\beta_{\mathcal{F}}$	$\alpha_{\mathcal{G}}$	$\beta_{\mathcal{G}}$
Australia	-18 ± 30	-12 ± 6	61 ± 2	6.2 ± 0.4
Belgium	510 ± 16	87 ± 3	138 ± 5	20 ± 1
Canada	-142 ± 20	-35 ± 4	83 ± 2	11.7 ± 0.4
France	356 ± 11	55 ± 2	43 ± 1	3.7 ± 0.2
Germany	549 ± 19	90 ± 4	55 ± 2	5.9 ± 0.4
Japan	581 ± 23	94 ± 4	13 ± 1	1.0 ± 0.2
Netherlands	504 ± 14	85 ± 3	78 ± 2	9.8 ± 0.4
Norway	603 ± 51	110 ± 10	45 ± 4	5.6 ± 0.8
Sweden	578 ± 35	93 ± 6	34 ± 1	3.2 ± 0.3
Switzerland	322 ± 22	49 ± 4	47 ± 1	5.9 ± 0.2
United Kingdom	273 ± 15	48 ± 3	69 ± 2	7.6 ± 0.4
United States	295 ± 14	54 ± 3	127 ± 3	17.7 ± 0.6

Table 5: All values are expressed in 10^{-4} units.

the linear functions \mathcal{F} and \mathcal{G} have been estimated, the next relevant parameter to fix is the rate of exponential growth of averaged earnings g . This can be easily obtained iteratively, first regressing a vector of 196 monthly logarithmic averages on a linear function of h , then rolling the window and repeating the estimate from January 1881 until the last admissible date. In this way we obtain a sample of g values from which empirical 68% confidence intervals can be computed, for the results on a monthly basis from Shiller's data series for the US market please refer to Table 7. The value of θ and σ_d are estimated regressing $\Delta(d_{t-1} - X_t)$ on $d_{t-1} - X_t - \log \mathcal{G}$, employing the same technique of rolling windows discussed above. The measured mean value for θ is equal to 0.0207, and following the consideration written at the end of the previous section this value implies a typical scale for mean-reversion of order 48 months. The parameters of the μ_t process are also estimated by means of linear regression of μ_{t+1} on μ_t , and on $\log \langle e \rangle_t^{10} - X_t + \mathcal{H} + g\mathcal{F}t$. In the first instance this require to introduce a suitable proxy for the unobservable variable μ_t . Our approach has been to estimate it by means of $X_t - X_{t-1}$, but different choices are admissible, for example in terms of an exponential weighted average of past price returns. However in this latter case an undesirable dependence of the parameters on the value of the exponential weight would be introduced. Moreover \mathcal{H} is still an unknown quantity that we have to choose before performing the optimization procedure. Recalling ex-

MSCI Belgium					MSCI France				
h	$\frac{1}{h}\mathbb{E}_0\left[\sum_{i=0}^{h-1}\log\left(1+\frac{D_i}{P_{i+1}}\right)\right]$	$\beta \pm \epsilon_\beta$	R^2	$\frac{g}{h}\mathbb{E}_0\left[\sum_{i=0}^{h-1}\left(1+\log\frac{D_i}{P_{i+1}g}\right)\right]$	$\frac{1}{h}\mathbb{E}_0\left[\sum_{i=0}^{h-1}\log\left(1+\frac{D_i}{P_{i+1}}\right)\right]$	$\beta \pm \epsilon_\beta$	R^2	$\frac{g}{h}\mathbb{E}_0\left[\sum_{i=0}^{h-1}\left(1+\log\frac{D_i}{P_{i+1}g}\right)\right]$	
	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	
24	222 ± 8	34.1 ± 1.5	63%	202 ± 6	31.0 ± 1.1	72%	65 ± 2	7.3 ± 0.4	
48	184 ± 7	27.5 ± 1.3	59%	173 ± 6	25.9 ± 1.1	67%	51 ± 2	4.8 ± 0.3	
72	153 ± 6	21.8 ± 1.2	57%	144 ± 5	20.6 ± 1.0	63%	45 ± 1	3.7 ± 0.3	
96	137 ± 5	18.8 ± 0.9	63%	128 ± 4	17.6 ± 0.8	69%	43 ± 1	3.5 ± 0.2	
120	124 ± 4	16.6 ± 0.8	67%	116 ± 3	15.4 ± 0.7	71%	43 ± 1	3.5 ± 0.2	
144	113 ± 4	14.6 ± 0.7	70%	103 ± 3	13.2 ± 0.6	72%	41 ± 1	3.3 ± 0.1	
168	100 ± 3	12.1 ± 0.7	66%	86 ± 3	10.0 ± 0.6	62%	38 ± 1	2.7 ± 0.1	
192	86 ± 3	9.5 ± 0.6	62%	67 ± 3	6.4 ± 0.6	43%	34 ± 1	1.9 ± 0.1	
MSCI United Kingdom					MSCI Germany				
h	$\frac{1}{h}\mathbb{E}_0\left[\sum_{i=0}^{h-1}\log\left(1+\frac{D_i}{P_{i+1}}\right)\right]$	$\beta \pm \epsilon_\beta$	R^2	$\frac{g}{h}\mathbb{E}_0\left[\sum_{i=0}^{h-1}\left(1+\log\frac{D_i}{P_{i+1}g}\right)\right]$	$\frac{1}{h}\mathbb{E}_0\left[\sum_{i=0}^{h-1}\log\left(1+\frac{D_i}{P_{i+1}}\right)\right]$	$\beta \pm \epsilon_\beta$	R^2	$\frac{g}{h}\mathbb{E}_0\left[\sum_{i=0}^{h-1}\left(1+\log\frac{D_i}{P_{i+1}g}\right)\right]$	
	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	R^2	$\alpha \pm \epsilon_\alpha$	$\beta \pm \epsilon_\beta$	
24	83 ± 3	9.8 ± 0.6	42%	83 ± 3	10.1 ± 0.6	45%	72 ± 3	8.5 ± 0.6	
48	75 ± 3	8.4 ± 0.6	38%	75 ± 3	8.8 ± 0.6	41%	62 ± 3	6.8 ± 0.5	
72	73 ± 3	8.0 ± 0.6	42%	73 ± 3	8.4 ± 0.6	44%	59 ± 2	6.3 ± 0.4	
96	75 ± 2	8.5 ± 0.5	55%	76 ± 2	8.9 ± 0.5	56%	60 ± 2	6.5 ± 0.4	
120	75 ± 2	8.7 ± 0.4	64%	76 ± 2	9.2 ± 0.4	65%	58 ± 2	6.3 ± 0.3	
144	72 ± 2	8.1 ± 0.4	70%	74 ± 2	8.8 ± 0.4	73%	54 ± 2	5.7 ± 0.3	
168	65 ± 1	6.8 ± 0.3	75%	66 ± 1	7.2 ± 0.3	77%	47 ± 2	4.5 ± 0.3	
192	56 ± 1	4.9 ± 0.3	70%	55 ± 1	4.9 ± 0.3	64%	41 ± 1	3.3 ± 0.2	

Table 6: All values for α , β , and associated errors are expressed in 10^{-4} units. Significance codes: 0 ‘’, 0.01 ‘*’, 0.1 ‘**’, 0.5 ‘***’.

$g (\times 10^{-4})$	$\theta (\times 10^{-4})$	γ	$\kappa (\times 10^{-4})$	$\sigma_d^2 (\times 10^{-4})$	$\sigma_\mu^2 (\times 10^{-4})$	$\sigma_X^2 (\times 10^{-4})$
12 (-3, 31)	207 (62, 352)	0.26 (0.19, 0.32)	201 (68, 323)	13 (11, 26)	12 (10, 21)	18 (17, 19)

Table 7: Estimated values of the parameters from Shiller’s time series and 68% confidence intervals.

pression (14), it is reasonable to assume it linear in the log EP ratio $\mathcal{H} = \alpha_{\mathcal{H}} + \beta_{\mathcal{H}} \log EP_0$ and to initialize the procedure with arbitrary values for the linear coefficients. The numerical process then produces optimal estimates of γ and κ , for suitable $\alpha_{\mathcal{H}}$ and $\beta_{\mathcal{H}}$ minimizing the leading correction term (14). In particular, for $\alpha_{\mathcal{H}} = 1.10$ and $\beta_{\mathcal{H}} = -0.79$ we have found $\gamma = 0.26$ and $\kappa = 0.02$ both satisfying the constraint $4\kappa \leq (1 - \gamma)^2$. For the estimate of σ_X we approach the problem from a different point of view. Indeed from equation (7) we know that the variance of $(X_h - X_0)/h$ scales as one over h with proportionality constant equal to σ_X^2 . Estimating the variance of price returns over h months for h ranging from 24 until 196, and fitting the curve with a linear relation we have found the value of 0.18% reported in Table 7. Ultimately, we test the goodness of the estimated parameters and in Figure 4 we present the results of a Monte Carlo simulation of the model (5)-(12). Each point corresponds to a single realization of $y_{t,h}$ with $h = 24, \dots, 196$ months with t starting from January 1881. The initial time values of Y_0 , μ_0 , and $\log DP_0$ are fixed equal to the empirical ones. We have also plotted the linear relation between yields and the log EP predictor with coefficients specified in Table 5. The numerical test confirms the ability of the proposed model to capture the relevant features of the historical time series, in particular the shrinking of the data cloud to the expected value with a scaling exponent dominated by the diffusive component of the price dynamics.

5 Conclusion and Perspectives

This paper documents substantial predictability of stock index returns on the basis of a simple valuation metric: the cyclically adjusted price-earnings ratio.

Practitioners and academicians alike have been using several valuation measures for estimating the intrinsic value of a stock index: for example in table 2 of [Poterba and Samwick \(1995\)](#) the ratio of market value of corporate stock to GDP, the year-end price-to-earnings ratio, the year-end price-to-dividend ratio and Tobin’s q are reported from 1947 to 1995 in an effort

of alerting the reader on the possible overvaluation of the index⁵. In particular Tobin’s q has been proposed as another efficient method of measuring the value of the stock market, with an efficiency comparable to the CAPE [Smithers \(2009\)](#). The q ratio is the ratio of price to net worth at replacement cost rather than the historic or book cost of companies. It therefore allows for the impact of inflation, much alike the CAPE which averages real earnings over a ten year span. It would be interesting to carry out an empirical analysis of the relationship between Tobin’s q and future stock index returns also for countries other than the U.S., but we have not been able to find high quality long-term time series of the q -ratio outside the U.S..

Substantial evidence of the importance of fundamentals in the valuation of international stock markets has been accumulated by the proponents of fundamental indexation [Arnott et al. \(2005\)](#).

Clearly our empirical analysis should be extended so as to cover longer historical periods. Longer time-series would also be important so as to investigate possible “evolutionary” phenomena like the end of the relationship between market valuation and interest rates, which may perhaps be interpreted as an example of the adaptive market hypothesis [Farmer and Lo \(1999\)](#), [Lo \(2004\)](#).

Acknowledgements

GB acknowledges support of the Scuola Normale Superiore Grant ‘Giovani Ricercatori 2011’.

References

Robert D. Arnott, Jason Hsu, and Philip Moore. Fundamental indexation. *Financial Analysts Journal*, 61(2):83–99, 2005.

⁵ “It is difficult to distill a simple conclusion from table 2. While price-to-earnings ratios are not unusually high at present, other measures of stock price valuation are at, or near, historical highs. (...) Table 2 does suggest, however, that in assessing the macroeconomic consequences of stock price movements, it may be important to distinguish between stock price fluctuations that are associated with movements in the price-to-earnings or price-to-dividends ratios and those that are not. A number of recent studies suggest that variations in the earnings-to-price ratio are correlated with prospective stock market returns, (...). Sharp increases in either the price-to-earnings or the price-to-dividends ratio, other things equal, are associated with lower prospective returns.”, see [Poterba and Samwick \(1995\)](#) pages 309–310.

- John Y. Campbell and Robert J. Shiller. Cointegration and tests of present value models. *Journal of Political Economy*, 95(5):1062–1088, 1987.
- John Y. Campbell and Robert J. Shiller. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, 1(3):195–228, 1988a.
- John Y. Campbell and Robert J. Shiller. Stock prices, earnings and expected dividends. *Journal of Finance*, XLIII(3):661–676, 1988b.
- John Y. Campbell and Robert J. Shiller. Valuation ratios and the long-run stock market outlook. *The Journal of Portfolio Management*, 24(2):11–26, 1998.
- Carl Chiarella. The dynamics of speculative behaviour. *Annals of Operations Research*, 37(1):101–123, 1992.
- Eugene F. Fama. The behavior of stock-market prices. *The Journal of Business*, 38:34–105, 1965.
- Eugene F. Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2):383–417, 1970.
- Doyne Farmer and Andrew W. Lo. Frontiers of finance: Evolution and efficient markets. *Proceedings of the National Academy of Sciences*, 96(18):9991–9992, 1999.
- Benjamin Graham and Spencer B. Meredith. *The Interpretation of Financial Statements: The Classic 1937 Edition*. HarperBusiness, 1998.
- Benjamin Graham, David L. Dodd, and Sidney Cottle. *Security Analysis*. McGraw-Hill, New York, 1962.
- Joel Lander, Athanasios Orphanides, and Martha Douvogiannis. Earnings forecasts and the predictability of stock returns: Evidence from trading the S&P. Technical report, Board of Governors of the Federal Reserve System, 1997.
- Andrew W. Lo. The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management*, 30(5):15–29, 2004.

- Benoit Mandelbrot. The variation of certain speculative prices. *The Journal of Business*, 36(4):394–419, 1965.
- James Myers and Bhaskaran Swaminathan. What is the intrinsic Value of the Dow? *Journal of Finance*, 54(5):1693–1741, 1999.
- James M. Poterba and Andrew A. Samwick. Stock ownership patterns, stock market fluctuations, and consumption. *Brookings Papers on Economic Activity*, 1995(2):295–372, 1995.
- Paul Samuelson. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(2):41–49, 1965.
- Robert J. Shiller. *Irrational Exuberance*. Princeton University Press, 2000.
- Andrew Smithers. *Wall Street Revalued. Imperfect Markets and Inept Central Bankers*. John Wiley and Sons, Ltd, 2009.
- Robert A. Weigand and Robert Irons. Does the market P/E ratio revert back to “average”? *Investment Management and Financial Innovations*, 3(3):30–39, 2006.
- John Burr Williams. *The Theory of Investment Value*. Harvard University Press, 1938.